Title: Effect Of Override Size On Forecast Value Add

By Name(s): Jeffrey A. Baker
Advisor: Dr. Sergio Caballero

Topic Areas: Demand Planning, Forecasting, Database Analytics

Summary: Business forecasting leverages qualitative overrides to create a final forecast. The objective is to reduce forecast error, which enables safety stock reduction, customer service improvement, and manufacturing schedule stability. However, overrides often fail to improve final forecast accuracy. This thesis offers a framework which leverages machine learning to identify non value added overrides. In turn, this can maximize the value that experts add to the business forecasting process.

Jeff is the founder and managing director of Libra Consulting. His expertise includes APS and S&OP, and he is proud to be in the very first blended cohort at MIT.

KEY INSIGHTS

1. **Judgmental overrides to forecasts often fail to increase forecast accuracy and represent a significant waste of effort.**
2. **Overrides are often entered to match the financial aspirations of the company and have a negative impact on forecast accuracy.**
3. **Classification methods can be utilized to guide forecasters as to whether a forecast is likely to add value.**

Introduction

Business forecasting frequently combines statistical time series techniques with qualitative expert opinion overrides to create a final consensus forecast. The objective of these overrides is to reduce forecast error, as measured by Mean Absolute Percent Error (MAPE). A common measure of the effectiveness of these overrides is Forecast Value Add (FVA), which simply measures the MAPE improvement due to the override. Lowering MAPE enables safety stock reduction, customer service improvement, and manufacturing schedule stability.

However, overrides often fail to improve final forecast accuracy. Process mis-steps include small adjustments, adjustments to accurate statistical forecasts, and adjustments to match financial goals. At best, these overrides waste scarce forecasting resources; at worst, they seriously impact business performance.

These impacts are highlighted in a case study of 703 overrides, which were split into 4 groups based on percentage size of the override (Figure 1). The smallest 25% of adjustments did little to improve forecast accuracy, and represent wasted effort. Larger downward adjustments tended to improve accuracy. In most business settings, downward revisions are thoroughly vetted, resulting in better decisions on the forecast.

Larger upward revisions, however, reduced accuracy. Upward revisions are often the result of pressure to hit financial goals. This creates a gap between the unbiased forecast and the aspirational financial goals. Without a plan to bridge that gap, forecast accuracy suffers.

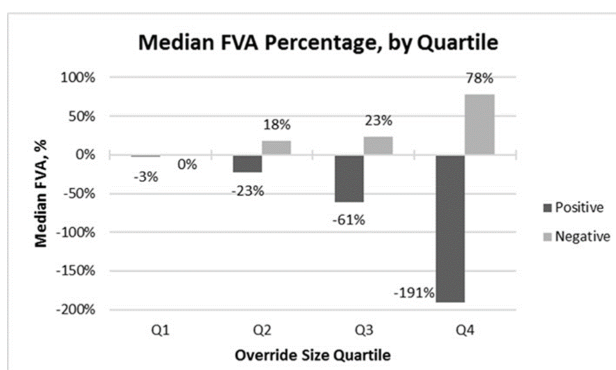

Figure 1: Median Forecast Value Added Percentages, based on Override Size Quartile.

Research Questions

The research question was whether a framework could be created to identify overrides which were likely to add value based on the FVA metric. Small adjustments would be ignored, and potentially value destroying overrides could be flagged and vetted.

Other questions revolved around the override size; Does the size of the override in relation to the underlying variability impact the ability to improve upon it? Does direction of forecast adjustment matter? And finally, does the baseline statistical forecast performance impact improvement opportunities?

Methodology

This thesis created a framework for identifying overrides that are likely to improve forecast accuracy, thereby increasing the efficiency and effectiveness of the forecasting process.

A new class of metrics, Dispersion-Scaled Overrides, was developed as an input to the framework. Other inputs included statistical forecast accuracy and auto-correlation.

To address this challenge, a framework was developed to classify proposed overrides as value added or non-values added based on two categories of predictor variables.

This approach is appropriate for any company using time series forecasting techniques and managerial overrides. It requires time series forecasting “triples”: actual demand, statistical forecast, and final consensus forecast. These triples must be collected at the same level where the override is entered, as that is where the business rationale behind the override is developed and challenged through the consensus process. Typically, the consensus process occurs during sales and operation planning, but this is not a requirement. If consistency is maintained, any level of the forecasting hierarchy may be used.

The first step in the methodology is to utilize the forecasting triples to create a predictor variable for each set of data points. The actual demand was used to determine the percentage errors of statistical and final consensus forecasts. Then, those two error values were compared to create an improvement metric known as Forecast Value Added (FVA). FVA reflects the success of the override in reducing forecast error.

FVA was then compared to a user-defined FVA threshold, FVA_{crit}. Values less than this value are classified as non-value added, while those above the threshold are classified as value added. This is the response variable of interest.

The predictor variables fell into two categories. The first was Dispersion-Scaled Overrides (DSO), a new metric created in this research. The product demand time series was decomposed into trend, seasonal, and residual components. For the residual component, the dispersion measures of standard deviation, trimmed mean absolute deviation, and median absolute deviation were calculated. The judgmental override was divided by these measures of dispersion, creating a signal-to-noise relationship between the override and the underlying variability in the time series. Different dispersion metrics were utilized due their varying responsiveness to outliers. Standard deviation is known to be impacted by outliers, whereas median absolute deviation is more robust to outliers.

The second category of predictor variables were opportunity indicators, which quantify where the statistical model may be under-performing, In these cases, expert intervention using overrides would likely reduce forecast error. One predictor value was autocorrelation in the residuals, which would indicate if there was information not captured in the statistical forecast. The others were percentage-based measures of statistical forecast accuracy – Root Mean Square Percentage Error (RMSPE), Mean Absolute Percent Error (MAPE) and Median Absolute Percent Error (MdAPE). Again, different measures were utilized based on their robustness or sensitivity to outliers.

The FVA response variable and predictor variables were then analyzed using the machine learning techniques described above to classify an override as value added.

The data mining techniques of classification tree, boosted tree, random forest, and logistic regression were chosen. These classification techniques have several advantages, including automatic variable selection, robustness to outliers, and no assumptions regarding linearity of relationships between predictor variables and the response variable. In particular, classification trees are attractive because they are graphical in nature and easy to explain to non-technical business personnel. Figure 2 outlines the methodology used in this work.

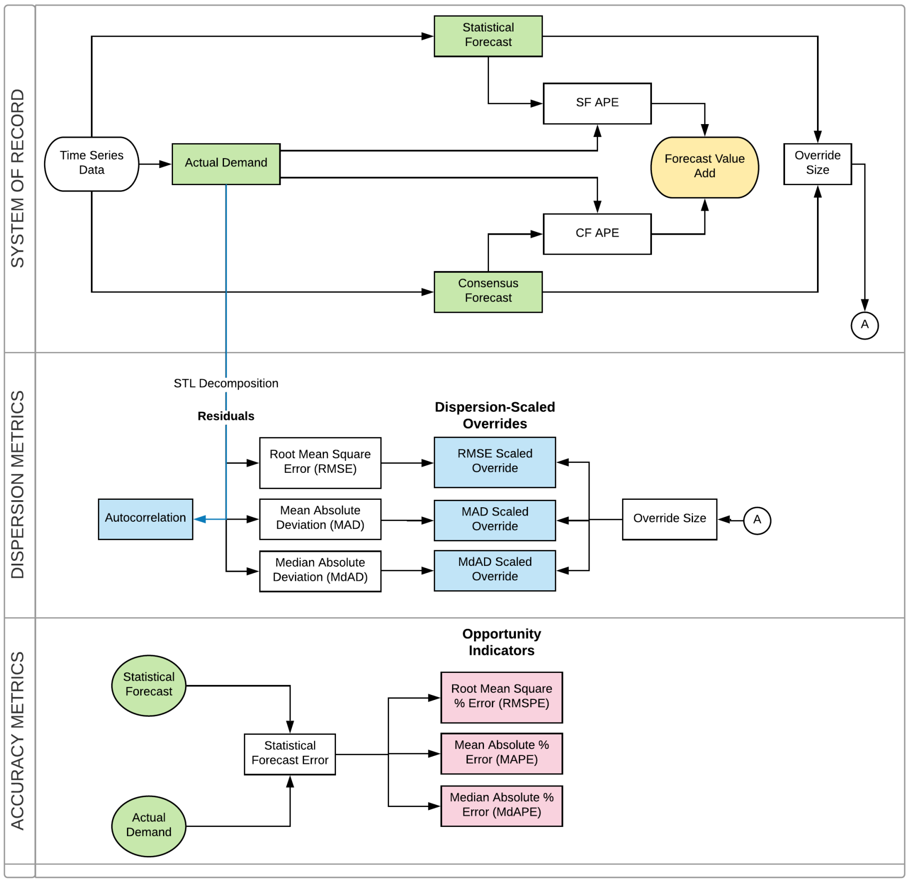

Figure 2: Methodology Overview

Results

The classification framework created was approximately 80% accurate in predicting whether an override would or would not create forecast value add above a user-defined threshold. This suggests that using Dispersion-Scaled Overrides alongside common forecast accuracy metrics can reliably predict forecast value add. In turn, this can maximize the value that experts add to the business forecasting process. Three key, interrelated factors drive the probability of creating a value-added override.

The first factor is the accuracy of the baseline statistical model. If the statistical forecast was performing poorly, it was more likely that there were opportunities to improve upon it. Conversely, a well-performing statistical model was difficult to improve upon.

The second factor was the size of the override scaled by the residual variation of the time series – a new class of metrics called Dispersion-Scaled Override (DSO). Small overrides which are indistinguishable from the underlying random noise variation are unlikely to add value. Larger overrides with a clear signal-to-noise have a better opportunity to improve accuracy.

The final factor was the direction of the override. Downward adjustments of the statistical forecast were more likely to add value, likely due to increased scrutiny. Upward adjustments, often driven by the desire to match financial goals, were more likely to degrade forecast accuracy. This directional bias is consistent with some previous studies and is a clear opportunity for improvement in the business forecasting process.

While autocorrelation was considered, it did not appear to play a statistically significant role in predicting forecast value add.

Conclusions

In this thesis, a classification framework for identifying value added overrides which improve forecast accuracy was created. This framework should be used during the consensus forecasting process to evaluate any override to the baseline statistical model.

In a case study at a sporting goods retailer, the classification framework was approximately 80% accurate in predicting success for the business studied in this research. Despite the presence of biased overrides, the framework demonstrated the capability to adjust.

The results suggest that using dispersion-scaled overrides alongside forecast accuracy metrics in a classification framework can reliably predict forecast value add. There are two key impacts to the business forecasting process which will be impacted.

First, numerous small adjustments which typically do not add value will be avoided. Additionally, adjustments to already accurate statistical forecasts will be contraindicated. Both will reduce the burden on the forecaster, allow them to concentrate on more important opportunities.

Second, the relationship captured by the DSO metric will drive cross-functional conversation and consensus during the sales and operations planning process. Initial overrides flagged as non value add require additional vetting. The forecast adjustment behavior will become self-correcting. Previous bad decisions will increase the likelihood of an override being flagged as non value add, which will increase the need to document and discuss underlying assumptions. Documentation and discussion will in turn increase forecast value-add.

This may be visualized in Figure 3.

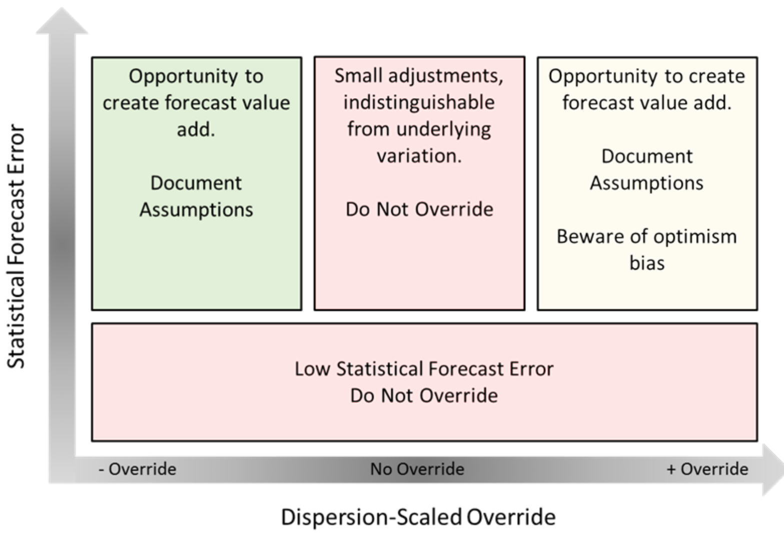

Figure 3: Relationship between Dispersion Scaled Overrides and Statistical Forecast Error

These two key advancements lay the groundwork for an efficient and effective business forecasting process. The impact will be seen in optimum inventory levels, increased manufacturing schedule stability, and improved customer service. Ultimately, business profitability will improve.